

## Berber subclassification

Maarten Kossmann  
Leiden University

Berber languages are a close-knit language group, whose internal differences remind one of those to be found inside Germanic or Romance. In Berber scientific and political discourse, there is a tendency to play down the differences, and often Berber is represented as one single language with only some superficial regional variation (e.g. Chaker 1995:9). On the other hand, some sources outside the world of Berberology use a very liberal definition of “language” where Berber is concerned; thus the *Ethnologue* (Lewis 2009) has no less than 25 different Berber languages.

Instead of providing a tree-model of the different varieties, I shall proceed by defining a number of “blocks”, i.e. bundles of varieties that have fairly great consistence.

Two such blocks are easily defined. In the first place, Zenaga in Mauritania stands on its own. It is in many points very different from other Berber languages, e.g. by showing entirely different developments in phonology, and several highly original morphological traits. Zenaga has one sister language, Tetserrét, which is spoken by parts of the Aytawari Seslem, a small sub-group of the Iwellemmeden Tuareg ethnicity in Niger, thousands of kilometers from Zenaga territory. As shown conclusively by Lux (2011), this language has many innovations in common with Zenaga and should be considered a sister variety of it.

The other clearly defined group is Tuareg, which has several regional variants in Mali, Niger and Algeria, and, to a lesser extent in Burkina Faso (due to immigrations from the 18<sup>th</sup> century onwards) and Libya. Again, this group is defined by a large number of common innovations. Internal diversification is important, and speakers from different varieties may have difficulties in understanding each other.

Zenaga, Tetserrét and Tuareg are the only Berber languages that fall outside the northern Berber dialect continuum that stretches from the Atlantic coast in Morocco to the Siwa oasis in Egypt. Although sometimes large Arabic-speaking territories lie between several patches of Berber speakers, adjacent (or the like) dialects are normally mutually understandable, and communication using Berber is possible. The effect of this continuum situation is that linguistic innovations do not hurt at linguistic obstacles; they can freely spread over the continuum, leading to an intricate network of isoglosses, which only rarely define clear linguistic boundaries.

Berber dialectology has mostly been strictly synchronic, trying to define different synchronic blocks, using different approaches (e.g. Willms 1980; Ameur 1990; Lafkioui 2009). While some minor breaks appear, most of the results confirm a basically geographic cline, not so different from the situation invoked by André Basset, the foremost Berber dialectologist of the 20<sup>th</sup> century: “ce n’est même pas, comme on le croit trop généralement, une langue divisée en quelques dialectes (...). Il en résulte que cette langue s’éparpille directement ou à peu près en une poussière de parlers, de 4 à 5 mille peut-être pour quelques cinq millions d’individus” (Basset 1952:1). [“it does not even constitute – as is too often believed – a language divided into a small number of dialects

(...). As a result, this language fragments, immediately or almost so, into a dust cloud of varieties, four to five thousand on a population of about five million people.”]

Different results are found when taking a historical point of view (e.g. Kossmann 1999, Naït-Zerrad 2001). This follows two itineraries. In the first place, certain innovations that are believed to be quite early in the development of Berber, and which are very commonly found, can be used in order to single out varieties that did not undergo this development (i.e. have archaic features where the others shared an innovation). In the second place, when a consistent bundle of unrelated isoglosses defines a territory, one may venture the idea that these varieties once formed a unity. The many much less consistent isoglosses cutting through these territories are then interpreted as later innovations.

Using these methods, a number of historically defined entities can be distinguished, which will be described following the northern Berber territory from west to east.

In the western part of Morocco, there is a large continuous territory which covers the Anti-Atlas, the Sous plains, the High Atlas and most of the Middle Atlas. The medieval materials contained in a number of texts (van den Boogert 1997, van den Boogert 2000) clearly belong to this group. While the varieties spoken at the extremes of the continuum are very different from each other, there are no clear internal boundaries inside it – not even the abrupt relief of the High Atlas chain seems to constitute a major break, and the dialects spoken north and south of it are quite similar. For practical reasons, often a distinction is made between two main varieties, Tashelhiyt, spoken in southwestern Morocco in the Anti Atlas, the Sous plains and the western High Atlas, and Tamazight (aka Middle Atlas Berber),<sup>1</sup> spoken in the southeastern Moroccan oases, in the eastern High Atlas and in the western and central Middle Atlas. While differentiating between these entities may be useful for practical purposes, and reflects the amount of linguistic variation, it is deceiving in suggesting a clear boundary between the two.

In northwestern Morocco, two more varieties are spoken that may belong to the same western Morocco block: Ghomara Berber and Senhadja de Sraïr.<sup>2</sup> It seems, however, that they have been without contact to the other western Moroccan varieties for a while, which has lead to many untypical retentions and innovations; a more detailed investigation might reveal that one of them (Ghomara), or both, are best considered entities on their own.

Eastern Morocco, most of Algeria, Tunisia and parts of Libya belong to one single historical block, which has been called Zenatic since the late 19<sup>th</sup> century (cf. for an early overview Destaing 1920).<sup>3</sup> This block is characterized by an important number of innovations in phonology and morphology (recently Kossmann 1999; Naït-Zerrad 2001). Taking these innovations as a basis, Zenatic has clear boundaries in Morocco and in

---

<sup>1</sup> Tamazight is the name of the Berber language in a large number of Berber varieties (a.o. Riffian and, with altered phonology, Tuareg); therefore its use for one specific variety is unlucky. The more neutral term “Middle Atlas Berber”, which is often used instead, is unlucky too, as it also comprises parts of the High Atlas mountains. “Central Moroccan Berber” seems to be the least problematic term, but has not been used yet by many authors.

<sup>2</sup> In spite of claims to the contrary (e.g. Lewis 2009), both varieties are well alive, cf. El Hannouche 2010; Mourigh in prep.; Lafkioui 2007.

<sup>3</sup> The name derives from the important historical entity of the *Zanāta*. Whether there is a link between the historical and the linguistic entity is an open question. Note that our delimitation is different from that found in, for instance, the Ethnologue (Lewis 2009).

Algeria. In Morocco, this boundary separates Ghomara and Senhadja de Sraïr (non-Zenatic) from Riffian (Zenatic); in the Middle Atlas, it separates Ayt Seghrouchen (Zenatic) from its non-Zenatic western neighbors. In northern Algeria, Zenatic comprises all varieties except Kabyle. Zenatic is also the language of the major oases in the northern Sahara, Figuig, Gourara, Mزاب and Ouargla. More to the east, the boundaries of Zenatic become blurred. The Tunisian dialects (as far as we know about them) and Zuara (Libya) are still classical Zenatic varieties. However, the oases of Sokna, Elfoqaha and Siwa share some features with Zenatic, but lack other features. Thus, while Zenatic has clear boundaries to the west and with Kabyle, it is more in a relation of continuum with eastern varieties.<sup>4</sup> In view of the large and – because of the expansion of Arabic – rather scattered geographical distribution, Zenatic has been cut up in some accounts into many different “languages” (e.g. Lewis 2009, which has 9 languages inside what is here considered to be Zenatic). This is rather arbitrary; in fact, there is sometimes remarkable mutual understanding over long distances. Thus speakers of Figuig Berber (eastern Morocco) travelling to Libya were astonished that they could understand Zuaran without a problem, while they would not understand Moroccan Tashelhiyt.

Kabyle seems to stand alone. It has important dialectal fragmentation, and especially the most eastern varieties (eastern part of “Petite Kabylie”) are very different from what is found elsewhere. In a number of cases, Kabyle has undergone similar innovations to the western Moroccan block. It is difficult to decide, at this point, whether this points to an earlier extension of this block towards Algeria (separated by the incursion of Zenatic), or whether they represent parallel developments.

The situation in Libya and Egypt is the most complicated in Berber dialectology; unfortunately, we lack good documentation of some of the key Libyan dialects (esp. Sokna, Elfoqaha and Awdjilah, some of them probably extinct now). As mentioned above, the fishing port of Zuara has a classical Zenatic dialect. The oasis dialects of Sokna and Elfoqaha seem to represent varieties close to Zenatic, but not quite part of it. Siwa (Egypt) is relatively similar to Sokna and Elfoqaha, but has undergone major innovations, esp. in verbal morphology. These innovations are so profound that one suspects that it has undergone a kind of simplification due to the presence of a large community of non-native speakers in the oasis – something well-known from the history of the place (Souag 2010:18).

The dialect of Djebel Nefusa in western Libya has a special position. On the one hand, it reveals a number of very archaic features, which place it outside Zenatic and its eastern continuation, e.g. the retention of a continuant pronunciation of \*β before a consonant (Kossmann 1999:114). Moreover, it has a number of developments shared with Ghadames (esp. forms such as Nefusa *ufəs*, Ghadames *ofəs* ‘hand’ instead of general Berber (*a*)*fus*). On the other hand, some of the defining Zenatic developments are also found in Nefusan, thus establishing a link with this block. One remarks that Nefusan texts are quite easily to process with a knowledge of a Zenatic Berber variety, while this is much more difficult with Ghadames texts. Maybe Nefusan is best viewed as a basically non-Zenatic dialect which, at a certain moment underwent very strong influence from neighboring Zenatic varieties. At this point the importance of the Ibadite network should not be underestimated (Brugnatelli 2008). The Ibadites constitute an early branching of

---

<sup>4</sup> This suggests that the propagation of the Zenatic dialects – whether by demic or merely by linguistic expansion – went from east to west.

the Islamic creed; in Northern Africa, they persist in the Djebel Nefusa and in the Zenatic-speaking communities of Mزاب, Djerba and Zuara. The continuous contacts between these brothers-in-creed may have well brought about a certain “zenatification” of Nefusan.

Two Libyan oasis dialects stand apart: Ghadames in western Libya and Awdjilah in eastern Libya. While they share a number of important archaisms, there is hardly any sign of common innovations, and they are best considered different entities. Both are very different from the other Libyan and Egyptian dialects, although Awdjilah has some recent innovations in common with nearby Siwa, esp. in syntax.

In a block-like classification of Berber languages, one has therefore the following historically defined entities:

1. Zenaga block (Zenaga of Mauritania, Tetserrét in Niger)
2. Tuareg block
- 3a. Western Moroccan block (SW Morocco, Central Morocco, i.e. Tashelhiyt and most of Tamazight)
- 3b. possibly including NW Moroccan Berber (Ghomara, Senhadja de Sraïr)
4. Zenatic block (Eastern Morocco, Western Algeria, Saharan oases, Tunisia, Zuara) extending towards the east with Sokna, Elfoqaha, Siwa
5. Kabyle (N Algeria), possibly linked to the western Moroccan block
6. Ghadames (Libya), probably to be linked to Djebel Nefusa (Libya)
7. Awdjilah (Libya)

It should be noted that some of the most salient dialectal phonetic developments in Berber cut across these groups, and seem to represent later innovations. This is the case of the lenition of stops (so-called spirantisation), which is found all over northern Morocco, northern Algeria and Tunisia, as well as in Zenaga, and which cuts both the western Moroccan block and the Zenatic block in two. It is also the case of the loss of the accentual system, which happened in all non-Tuareg varieties of Algeria and Morocco, and of the reduction of the system of short vowels, which is found all over the northern continuum, except for Ghadames and possibly Awdjilah (Prasse 1989).

Moreover, one remarks the existence of some salient features with a highly erratic distribution, cutting across many of the above blocks. One example is the distribution of the pronunciation *t̪* instead of *d̪* (e.g. *aṭar* ‘foot’ vs. *aḍar*). While *d̪* is found throughout the Berber territory, including Zenaga and Tuareg, *t̪* is found in an arbitrarily looking scattering of dialects in the east (Siwa, Awdjilah, Nefusa), in Algeria (eastern Kabyle), in the Zenatic block (Ayt Warayn in the eastern Middle Atlas) and in the western Moroccan block (Dades region and Ghomara). Similarly the 2S subject marking on the verb has three variants with a discontinuous distribution. Most generally attested is *t-...-əd*. However, a number of dialects have a pharyngealized consonant in the suffix (*t-...-əḍ* or

*t*...-*ə*), e.g. Siwa and Kabyle, while others have a non-pharyngealized voiceless suffix (*t*...-(*ə*)*t*), e.g. Awdjilah, Ghadames, Ghomara and Tashelhiyt. This suggests that at earlier times, before the emergence of the currently definable blocks, other entities existed, and that the groups defined above include members of many different earlier groups.

As a consequence, one may doubt whether the tree model is suitable for the description of the Berber language family. Its continuous history of convergence and differentiation along new lines makes any definition of branches arbitrary. Moreover, mutual intelligibility and mutual influence render notions such as “split” or “branching” rather difficult to apply, except, maybe, in the case of Zenaga and Tuareg.

Still, a number of elaborate attempts at subclassification have been made, using lexicostatistical methods.<sup>5</sup> In table (1) I present the results of one of these, Blažek’s (2010) tree based on 22 varieties using Starostin’s “calibrated” glottochronological method. The results are not too different from the seven-block compartmentalization presented above; the main difference is the place of Djebel Nefusa. Still, even if one accepts the basic tenets of the method, its application to Berber is difficult. The method discards loanwords from the set of items to be measured. This is relatively easy in the case of loanwords from Arabic. However, when it comes to borrowing between Berber varieties, it is hardly possible to distinguish loans from common heritage, both practically (how to see the difference) and theoretically (is a spreading lexical innovation to be considered common heritage or borrowing).

The dating of Proto-Berber is a difficult affair. As mentioned above, linguistic differentiation seems to be similar to that in Germanic or Romance, which would put it somewhere in the first millennium BC (cf. Louali & Philippon 2004). A similar date is provided by Blažek (680 BC). Other researchers have provided much earlier dates. Preliminary results from the Automated Similarity Judgment Program put the first branching (Zenaga) around 2000 BC (Müller e.a. 2010); in general, the datings for Berber in this analysis are extremely early; thus the differentiation between the neighboring dialects Beni Iznasen and Guelaya Riffian in Northern Morocco would lie about the year 300 AD, even though the dialects are mutually comprehensible and mainly differ in a few regular sound changes. Blench (2001) has proposed an even earlier date for Proto-Berber, around 4500 BC, explaining the high degree of uniformity from “highly mobile populations already speaking closely related languages, constantly encountering one another in open terrain” (184). Whatever the merits of the model as such, it is hardly applicable to northern Africa, which is for a large part mountain area.

In view of the continuous movement of convergence, one may ask whether the establishment of a Proto-Berber entity is possible at all (cf. Galand 2010:14). Múrcia Sánchez (2010:1104) has suggested that at a certain period, which he equals with proto-Berber, Berber would have constituted a koine rather than a unity, i.e. large-scale convergence would have blurred the distinctions between originally much more different

---

<sup>5</sup> An unconvincing classification is provided by Aikhenvald & Militarev (1991), cited and criticized by Blažek (2010). At many points this classification seems to be arbitrary; moreover at points it classifies dialects which are fully undocumented (e.g. Tmessa in Libya), or which are not Berber at all (e.g. Tadaksahak, which is Northern Songhay, and the Kufra oases, which are probably Teda-speaking). Unfortunately, some of the main lines of this classification have been taken over by Ethnologue (Lewis 2009).

varieties, without obliterating them entirely. Maybe even a more recent dating could be endeavored. One of the remarkable facts in Berber in antique sources (Múrcia Sánchez 2010) is the general use of *p* rather than *f* and *c* rather than *γ*. In modern Berber varieties, only *f* is found, and (non-geminated) *γ* is never pronounced as a velar plosive. Does this mean that in late Antiquity (or later?) these phonological innovations were able to spread all over the Berber speaking territory? Were these innovations accompanied by other innovations, less visible in the sources? To what extent do our reconstructions in fact represent this late Antique koineization rather than Proto-Berber? If this is the case, one may doubt that one might be able to get even a glimpse of the proto-language through the veil of the grand convergence movement(s) which followed it.

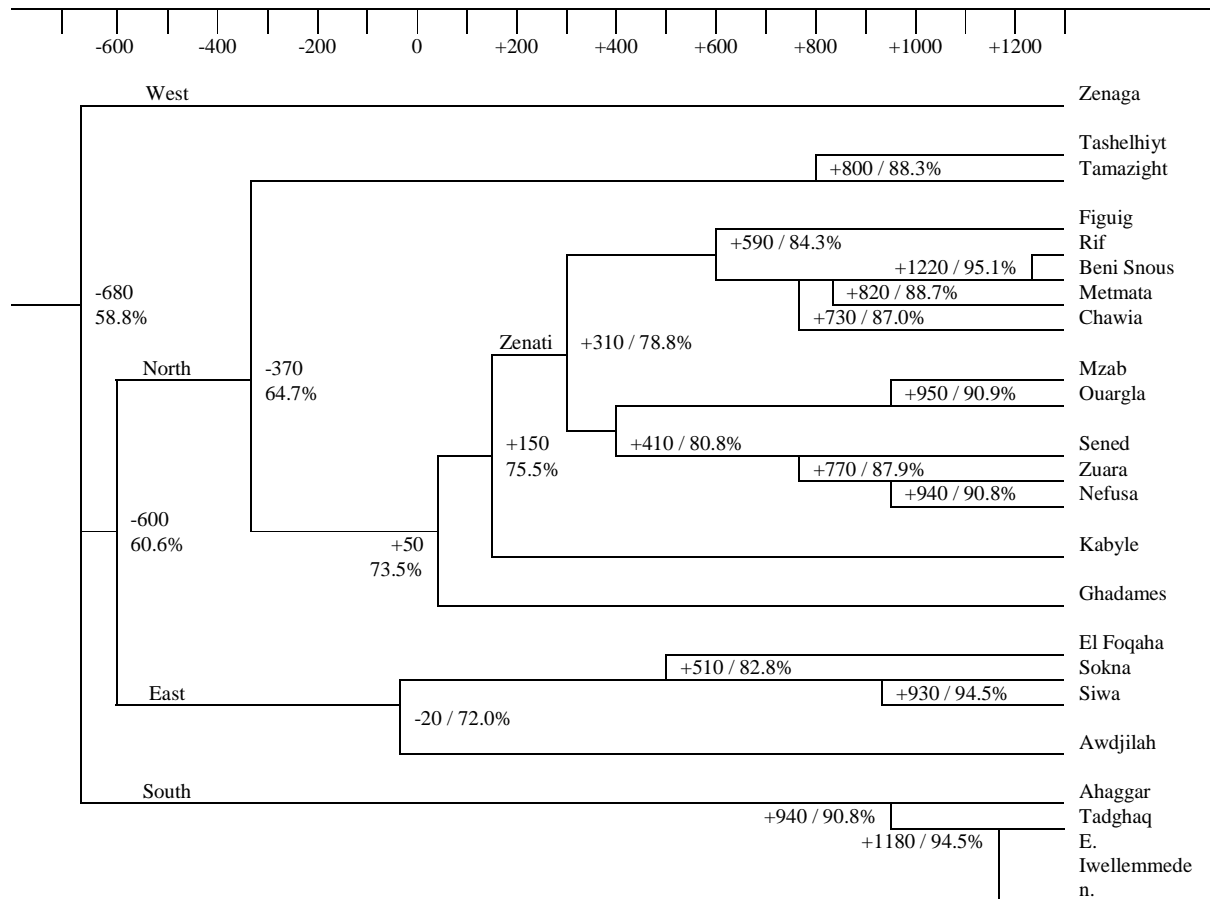
Two entities fall outside the normal definition of Berber in its synchronic sense, but clearly have some relationship with it. In the first place, somewhere in the first millennium BC writings appear in northern Africa, which use an alphabetic script, called “Libyan”, which seems to be to a large degree a proper invention. Kerr (2010) points to important structural parallels with the Punic script and posits its invention in the second century BC. Most specialists prefer an earlier dating (e.g. Pichler 2007), and it is conceivable that the arguments of Kerr rather concern an orthographic reform than the earliest design of the script. Although there are thousands of inscriptions in this script, their language is not easy to determine, as they almost exclusively consist of personal names. The few inscriptions that have somewhat more text show a language which has clear parallels with Berber; however, it is difficult to define this relationship with any precision (cf. Galand 2010:15-19).

The other entity is the ancient language of the Canary Islands,<sup>6</sup> commonly known as Guanche. This language died out somewhere in the 17<sup>th</sup> century, and we only have limited resources on it, which are more often than not difficult to interpret (cf. Wölfel 1965). There exists no doubt that the language contains Berber elements; there are many words that have clear and unproblematic cognates in Berber, such as *ilfe* ‘pig’, cf. the generally attested northern Berber form *iləf* ‘pig’. On the other hand, one is struck by the presence of a large stock of vocabulary that does not have any resemblance to Berber whatsoever. Moreover, as remarked by Galand more than once (e.g. 2010:2-4), the few short texts desist any interpretation from a Berber point of view, and no Berber inflectional elements seem to appear in them. As to the lexicon, one remarks that most of the Berber materials concern agricultural terms, pertaining to crops, livestock and related concepts. Most (but not all) terms referring to basic concepts do not have a clear correlate in Berber. One could explain this by positing a double layering in the language: it would be basically non-Berber, but due to the assimilation of a later influx of Berber speakers, who may have introduced new agricultural practices and livestock, large numbers of Berber words entered the lexicon. Such an explanation is tempting, but must remain speculation because of the scarcity of documentation; thus we are much better informed about cultural lexicon than about lexicon concerning body parts or basic verbs.

---

<sup>6</sup> I should like to thank Marijn van Putten (Leiden University) for his help in this matter.

Table 1. Blažek's (2010) classification of Berber based on Starostin's calibrated glottochronology based on minimal values [language names slightly adapted]



## References

- Aikhenvald, Alexandra Y. and Alexander Y. Militarev. 1991. Ливийско-гуанчские языки [Lybia-Guanche languages]. In Igor M. Diakonoff & G. Sh. Sharbatov (eds.), *Языки Азии и Африки [Languages of Asia and Africa]*, Vol. 4.2: Афризийские языки [Afro-Asiatic languages], 148-267. Moscow: Nauka. [not consulted]
- Ameur, Meftaha. 1990. A propos de la classification des dialectes berbères. *Études et Documents Berbères* 7. 15-27.
- Basset, André. 1952. *La langue berbère*. London/New York/Toronto: Oxford University Press.
- Blažek, Václav. 2010. On classification of Berber. Paper presented at the 40th *Colloquium of African languages and Linguistics*. Leiden, August 23-25.
- Blench, Roger. 2001. Types of Language Spread and their Archaeological Correlates: The Example of Berber. *Origini* 23. 169-190.
- Boogert, Nico van den. 1997. *The Berber Literary Tradition of the Sous*. Leiden: Nederlands Instituut voor het Nabije Oosten.
- Boogert, Nico van den. 2000. Medieval Berber orthography. *Études berbères et chamito-sémitiques. Mélanges offerts à Karl-G. Prasse*, ed. by Salem Chaker and Andrzej Zaborski. Paris and Louvain: Peeters. 357-377.
- Brugnatelli, Vermondo. 2008. D'une langue de contact entre berbères ibadites. *Berber in Contact. Linguistic and Sociolinguistic Perspectives*, ed. by Mena Lafkioui and Vermondo Brugnatelli. Köln: Rüdiger Köppe. 39-52.
- Chaker, Salem. 1995. *Linguistique berbère. Etudes de syntaxe et de diachronie*. Paris and Louvain: Peeters.
- Destaing, Edmond. 1920. Note sur la conjugaison des verbes de formes C<sup>1</sup>eC<sup>2</sup>. *Mémoires de la Société de Linguistique de Paris* 22. 139-148.
- El Hannouche, Jamal. 2010. Arabic Influence in Ghomara Berber. MA Thesis, Leiden University. [http://www.alfa-desk.nl/ghomara/Arabic\\_Influence\\_in\\_Ghomara\\_Berber\\_by\\_J\\_el\\_Hannouche.pdf](http://www.alfa-desk.nl/ghomara/Arabic_Influence_in_Ghomara_Berber_by_J_el_Hannouche.pdf)
- Galand, Lionel. 2010. *Regards sur le berbère*. Milano: Centro Studi Camito-Semiti.
- Kerr, Robert. 2010. Some thoughts on the origins of the Libyco-Berber alphabet. *Études berbères V. Essais sur des variations dialectales et autres articles*, ed. by Harry Stroome, Maarten Kossmann, Dymitr Ibrizimow and Rainer Vossen. Köln: Rüdiger Köppe. 41-68.
- Kossmann, Maarten. 1999. *Essai sur la phonologie du proto-berbère*. Köln: Rüdiger Köppe.
- Lafkioui, Mena. 2007. *Atlas linguistique des variétés berbères du Rif*. Köln: Rüdiger Köppe.
- Lafkioui, Mena. 2009. Analyses dialectométriques du lexique berbère du Rif. *Études berbères IV*, ed. by Rainer Vossen, Dymitr Ibrizimow and Harry Stroome. Köln: Rüdiger Köppe. 133-150.
- Lewis, M. Paul (ed.), 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. <http://www.ethnologue.com/>. 2009
- Louali, Naïma and Gérard Philippon. 2004. Berber expansion into and within north-west Africa: a linguistic contribution. *Afrika und Übersee* 87. 105-130.



- Lux, Cécile. 2011. Etude descriptive et comparative d'une langue menacée : le tetserret, langue berbère du Niger. PhD Thesis. Université Lumière Lyon II.
- Mourigh, Khalid. In prep. A Grammar of Ghomara Berber. PhD Thesis. Universiteit Leiden.
- Müller, André, Søren Wichmann, Viveka Velupillai, Cecil H. Brown, Pamela Brown, Sebastian Sauppe, Eric W. Holman, Dik Bakker, Johann-Mattis List, Dmitri Egorov, Oleg Belyaev, Robert Mailhammer, Matthias Urban, Helen Geyer, and Anthony Grant. 2010. ASJP World Language Tree of Lexical Similarity: Version 3 (July 2010). <http://email.eva.mpg.de/~wichmann/WorldLanguageTree-003.pdf>
- Múrcia Sánchez, Carles. 2010. La llengua amaziga a l'antiguitat a partir de les fonts gregues i llatines. PhD Thesis. Universitat de Barcelona. <http://84.88.13.203:8080/handle/10803/1686>
- Naït-Zerrad, Kamal. 2001. Esquisse d'une classification linguistique des parlers berbères. *Al-Andalus-Maghreb* 8-9. 391-404.
- Pichler, Werner. 2007. *Origin and Development of the Libyco-Berber Script*. Köln : Rüdiger Köppe.
- Prasse, Karl-G. 1989. Awjili (parler berbère d'Augila). *Encyclopédie Berbère* 7. 1052-1055.
- Souag, Lameen. 2010. Grammatical Contact in the Sahara: Arabic, Berber, and Songhay in Tabelbala and Siwa. PhD Thesis. SOAS, London.
- Willms, Alfred. 1980. *Die dialektalische Differenzierung des Berberischen*. Berlin: Dietrich Reimer.
- Wölfel, Dominik Josef. 1965. *Monumenta Linguae Canariae. Die kanarischen Sprachdenkmäler: Eine Studie zur Vor- und Frühgeschichte Weißafrikas*. Graz: Akademische Druck- und Verlagsanstalt.